"Efficiency Space" –
A Framework for Evaluating
Joint Evaporation and Runoff Behavior

Randal Koster

Global Modeling and Assimilation Office, NASA/GSFC, Greenbelt, MD

Corresponding author address:

Randal Koster

Global Modeling and Assimilation Office

Code 610.1, NASA/GSFC

Greenbelt, MD 20771

randal.d.koster@nasa.gov

Abstract

At the land surface, higher soil moisture levels generally lead to both increased evaporation for a given amount of incoming radiation (increased "evaporation efficiency") and increased runoff for a given amount of precipitation (increased "runoff efficiency"). Evaporation efficiency and runoff efficiency can thus be said to vary with each other, motivating the development of a unique hydroclimatic analysis framework. Using a simple water balance model fitted, in different experiments, with a wide variety of functional forms for evaporation and runoff efficiency, we transform net radiation and precipitation fields into fields of streamflow that can be directly evaluated against observations. The optimal combination of the functional forms – the combination that produces the most skillful streamflow simulations – provides an indication for how evaporation and runoff efficiencies vary with each other in nature, a relationship that can be said to define the overall character of land surface hydrological processes, at least to first order. The inferred optimal relationship is represented herein as a curve in "efficiency space" and should be valuable for the evaluation and development of GCM-based land surface models, which by this measure are often found to be suboptimal.

Capsule Summary: The relationship between evaporation and runoff at the land surface is captured in a simple "efficiency space" framework.

1. Introduction: Efficiency Space

A land surface model, or LSM, is an essential component of a climate modeling system. The numerous studies that have quantified the impact of land processes on simulated climate (e.g., Shukla and Mintz 1982, Delworth and Manabe 1989, Dirmeyer 2000, Koster et al. 2000, Seneviratne et al. 2006, to name a few; see Seneviratne et al. 2010 for a review) underscore the importance of realistic treatments of land processes in climate models. Recognition of this importance has spawned multi-national and multi-institutional LSM evaluation projects such as the Project for the Intercomparison of Land-surface Parameterization Schemes, or PILPS (Henderson-Sellers et al. 1993, Chen et al. 1997, Wood et al 1998, Bowling et al. 2003), the Global Soil Wetness Project (Dirmeyer et al. 1999, Boone et al. 2004, Dirmeyer et al. 2006), and the more recent benchmarking project PLUMBER (PALS Land Surface Model Benchmarking Evaluation Project, Best et al. 2014).

One of the findings from PILPS is that in terms of LSM behavior and performance, the evaporation and runoff formulations of an LSM are inextricably linked (Koster and Milly 1997). (Note that throughout this paper, the term 'evaporation' is used in place of the more cumbersome 'evapotranspiration' to encompass all evaporation processes.) A land surface scheme's treatment of the evaporation process has a profound impact on the runoff it generates; similarly, the scheme's treatment of the runoff process has a first-order impact on evaporation. Perhaps counter-intuitively, an LSM with an excellent evaporation formulation will, when forced with realistic meteorology, produce poor evaporation rates if the model's runoff formulation is poor, and it will accordingly lead to biases in any atmospheric model that is connected to it. To produce a reliable treatment of land surface processes in an Earth system model, it is arguably critical to understand the evaporation-runoff connection.

65 The framework introduced herein can serve as a tool for investigating this connection.
66 Building on the work of Koster and Milly (1997) and Koster and Mahanama (2012), the
67 framework captures the essence of the evaporation-runoff connection in a very simple and
68 concise way: as a curve lying in "efficiency space".

69 In essence, the efficiency space framework is built on the assumption that evaporation
70 efficiency (the ratio of latent heat flux, λE, to net radiation, R_{net}, where λ is the latent heat of
71 vaporization) is a function of soil moisture, W:

72 $$\lambda E / R_{net} = \beta (W), \quad (1)$$

73 and on the further assumption that runoff efficiency (the ratio of runoff production, Q, over some
74 time period to precipitation, P) is also a function of soil moisture:

75 $$Q / P = F (W). \quad (2)$$

76 Under these two assumptions, the two efficiencies (i.e., $\lambda E/R_{net}$ and Q/P) can be related to each
77 other, as illustrated in Figure 1a. The red curve on the right shows an arbitrary functional
78 relationship between $\lambda E/R_{net}$ and W, and the blue curve directly below it shows a similarly
79 arbitrary relationship between Q/P and W. Note that for mathematical tractability, both
80 evaporation efficiency and runoff efficiency are assumed throughout this paper to increase
81 monotonically with soil moisture, reflecting the reasonable idea that wetter soils allow an easier
82 production of both evaporation and runoff. To capture the well-known plateau for wetter soils
83 associated with energy-limited evaporation (e.g., Manabe, 1969), the rate of increase at high soil
84 moisture can be set very low, as in the figure.

85 For a soil moisture of 0.3 (degree of saturation), these functions produce an evaporation
86 efficiency of about 0.3 and a runoff efficiency of about 0.05. For a soil moisture of 0.5, the
87 evaporation and runoff efficiencies are about 0.5 and 0.1, respectively. These pairings of

88 evaporation and runoff efficiencies, along with every other possible pairing determined from the
89 two functions, are plotted in "efficiency space" on the left side of Figure 1a. The efficiency
90 space plot essentially shows how evaporation and runoff efficiencies vary with each other – how
91 efficient, for example, the land surface is at converting precipitation into runoff when it is able to
92 convert, say, 40% of net incoming radiative energy into evaporation, and how the runoff
93 efficiency changes when the evaporation efficiency is reduced to 30% through the drying of the
94 soil.

95 Again, the evaporation and runoff efficiency functions on the right in Figure 1a are
96 arbitrary; a different set of functions would lead to a different curve in efficiency space. Figure
97 1b shows another possible example. The number of possible efficiency function combinations,
98 of course, is infinite. The corresponding Q/P-vs-$\lambda E/R_{net}$ curves could lie anywhere in efficiency
99 space, subject to the condition that Q/P and $\lambda E/R_{net}$ increase monotonically with each other.

100 This paper has two overarching goals: (i) demonstrating that the identification of a curve
101 in efficiency space is tantamount to characterizing overall hydrological behavior, and (ii) finding
102 the curve that best represents nature and thus best serves as a target for LSM development. It is
103 important to emphasize here that simple functional forms such as those on the right in Figure 1a
104 are, at best, crude first order approximations to the much more complex behavior seen in land
105 surface models and (presumably) in nature itself. Nevertheless, the functional forms capture
106 enough of the underlying controls of soil moisture on evaporation and runoff to make the curve
107 on the left meaningful and to allow nature's efficiency space curve, however it looks, to embody
108 in a valuable and concise way an essential aspect of surface hydrology.

109 Given the difficulty of measuring evaporation and soil moisture at the large scale, the
110 efficiency functions underlying nature's efficiency space curve, and thus the curve itself, cannot

111 be derived directly from observations. This paper thus employs an indirect approach to search
112 for nature's curve: the use of a simple water balance model in conjunction with streamflow,
113 precipitation, and radiation observations. The approach and the main results are presented in
114 Section 2. The paper continues in Section 3 with a discussion of the efficiency space framework
115 in the context of Budyko's (1974) landmark hydroclimatic analysis, and Section 4 discusses the
116 use of the framework for LSM development. The summary section (Section 5) includes a
117 discussion of the framework's use in the context of the new soil moisture data expected from
118 recent and upcoming satellite missions.

119 Underlying all of the discussions in the paper is the fundamental idea that it is the *joint*
120 response of evaporation and runoff production to variations in soil moisture that underlies
121 hydrological behavior. Focusing inordinately on one or the other in LSM development is likely
122 to lead to inadequate LSM performance.

123

124 **2. Analysis of Efficiency Space with a Simple Water Balance Model**

125

126 a. Model Structure

127 The tool used in this paper to evaluate different curves in efficiency space is the simple
128 water balance model (WBM) of Koster and Mahanama (2012). The essence of the model is
129 illustrated in Figure 2. In addition to a soil water holding capacity, the user chooses the model's
130 functional efficiency relationships [i.e., $\lambda E/R_{net} = \beta(W)$; $Q/P = F(W)$]; shown in the figure are
131 arbitrary sample functions. The model is then driven with observations-based daily precipitation
132 and net radiation forcing over a long period of time, typically several decades. At a given time
133 step, the WBM's $\beta(W)$ function, in conjunction with the net radiation input and the current soil

moisture state, determines the evaporation for the time step. Similarly, the WBM's F(W) function, in conjunction with the precipitation, determines the amount of runoff generated. The WBM's soil moisture state is updated based on the prescribed precipitation rate and the computed evaporation and runoff values. Precipitation is added to a snowpack reservoir during subfreezing periods; snowpack water melts and is added to the soil (essentially as precipitation) as the prescribed temperature warms.

This WBM is undeniably much simpler than the land surface models (LSMs) typically used in climate modeling systems. The WBM lacks temperature prognostic variables and associated energy balance calculations, and it includes no specific treatment of (for example) baseflow and interception loss. Also, unless otherwise stated, most applications of the WBM here utilize the same efficiency functions in every season of the year and across large continental domains (i.e., without regard to spatial variations in topography, soil texture, vegetation type, and other features that may affect efficiency – shown in some studies to be important [e.g., Yildiz and Barros [2007]). This simplicity, while inappropriate for an LSM working within an Earth system model, is nevertheless justified for the WBM given that it is designed to contain *only the most important, first-order controls* on evaporation and runoff production contained within a more complex LSM (generally implicitly in the latter – the net result of complex interacting parameterizations), thereby allowing the isolated study of these controls. The numerous features not included in the WBM are thus assumed, when acting in a full LSM, to induce mainly second-order modifications to the fluxes. While the appropriateness of such an assumption can be argued, and while more complex versions of the WBM can easily be envisioned, it must be remembered that each addition of complexity brings the WBM closer to a full LSM and thereby makes the WBM results more difficult to interpret. The WBM, despite its

157 simplicity, has been found in tests to reproduce successfully the first-order behavior of complex,
158 state-of-the-art LSMs. (See Koster and Mahanama (2012) for one example.)

159

160 b. Simulation Procedure

161 Following precisely the procedure of Koster and Mahanama (2012), the WBM is run on a
162 2.5°× 2.5° grid across the conterminous United States (CONUS) over the period 1948-2000. The
163 interannually-varying precipitation rates used to force the WBM are from the dataset of
164 Andreidis et al. (2005), and the net radiation forcing used consists of climatological seasonal
165 cycles derived from the Surface Radiation Budget (SRB) dataset (Gupta et al. 2006). The final
166 52 years of simulated streamflow data are evaluated against observed streamflow data, which
167 consist of naturalized streamgauge measurements in several large hydrological basins (see
168 Mahanama et al. [2012] for details).

169 Figure 3a illustrates a sample simulation. The red and blue curves in the left panel show,
170 respectively, the evaporation efficiency and runoff efficiency functions prescribed in the WBM.
171 The WBM is driven across CONUS for 53 years, and the runoff rates produced in the grid cells
172 lying within the Upper Mississippi Basin (upstream of the Grafton stream gauge site) are
173 combined into a basin-average annual rate for each simulation year. These simulated annual
174 rates are compared to observed streamflows in the rightmost panel of Figure 3a. This particular
175 combination of efficiency functions is seen to lead to significantly underestimated streamflow.

176 Figure 3b shows a second sample simulation. The efficiency functions applied here are
177 clearly different, particularly that for evaporation efficiency – in this second simulation, the
178 ability of the WBM to convert net radiation energy to evaporation is, in general, reduced. As a

179 result, the simulated streamflows for the Upper Mississippi have increased and are now, in fact,
180 too high.

181

182 c. Collapse of Functions onto Efficiency Space

183 Figure 4 illustrates a finding that underlies much of the analysis in this paper. Three
184 pairings of efficiency functions are provided in the top row of the figure. At first glance the
185 pairings look fundamentally different, and yet a closer look shows that all three map onto the
186 same single curve in efficiency space (lower left panel). Furthermore, all three pairings, when
187 implemented into the WBM, produce essentially the same hydrological behavior – the lower
188 right panel of Figure 4 shows that the time series of runoff for the three cases is essentially the
189 same. This example is, in fact, representative. In general, pairings of efficiency functions that
190 produce the same curve in efficiency space generate essentially the same hydrological fluxes in
191 the WBM.

192 This result can be considered in the context of the water balance equation:

193 $$P = E + Q + C_w \Delta W/\Delta t, \quad (3)$$

194 where C_w is the water holding capacity of the WBM and Δt is the time period considered. Using
195 (1) and (2), this can be written as

196 $$P = E + P\,G[\lambda E/R_{net}] + C_w \Delta W/\Delta t, \quad (4)$$

197 where G is the functional relationship between $\beta(W)$ and $F(W)$ in efficiency space. Now
198 consider the idealized limit of zero water holding capacity. In this case, given that P and R_{net} are
199 inputs, the equation collapses into one that defines a unique value of E for a time step and,

accordingly, also a unique value for Q. In other words, regardless of the evaporation function employed – regardless, for example, of which red curve in the top row of Figure 4 is used – a given curve in efficiency space (i.e., a given $G[\lambda E/R_{net}]$ function) implies a single set of values for evaporation and runoff and thus the same overall hydrological behavior.

Naturally, the water holding capacity term is not necessarily negligible, and this explains the slight differences seen in the simulated streamflow curves in Figure 4. Again, though, the overall strong agreement between the streamflow curves in this example is found to be standard through a comprehensive series of tests – for practical purposes, pairings of functions that map to the same efficiency space curve are found to be hydrologically very similar, supporting the use of efficiency space curves as an efficient means of characterizing the joint behavior of evaporation and runoff functions.

d. Efficiency Space Curves

i. Overall ranking. The efficiency space curve in the lower left panel of Figure 4, regardless of the pairing of efficiency functions that produced it, leads to streamflow values that are too low relative to observations (lower right panel). This efficiency space curve is reproduced in Figure 5a as the yellow curve, with the yellow color chosen to reflect this level of streamflow simulation skill. (Although no actual skill numbers are shown here, as this discussion focuses only on the relative skill levels associated with different curves, the skill is quantified with a root-mean-square error, or RMSE, calculation against observed ratios of annual streamflow to annual precipitation in the basin.) A pairing of efficiency functions that leads to the blue curve in Figure 5a produces WBM streamflow values that are even more inaccurate (this is indeed the

222 pairing underlying Figure 3b), whereas a pairing that leads to the red curve in Figure 5a produces
223 a relatively realistic simulation.

224 Using this color-coding scheme for streamflow simulation accuracy, and plotting the
225 higher scoring curves on top of the lower scoring ones, Figure 5b shows the results of testing
226 more than 23000 curves in efficiency space. Figure 5b (and corresponding plots for other basins)
227 holds a substantial amount of information and is indeed a centerpiece result of this paper. It
228 shows, for example, that any curve passing through the northwest quadrant of efficiency space
229 will produce a poor WBM simulation of streamflow. Most importantly, given that observations
230 are used to evaluate the skill levels, it provides an indication of the curve that effectively
231 operates in nature across the Upper Mississippi Basin. The darkest red curves correspond to the
232 highest skill values; any dark red curve will produce an accurate simulation and is thus a
233 reasonable estimate of nature's curve.

234 The highest scoring curve (again, not by a large margin) is highlighted with a heavy black
235 line. Assuming that this particular curve reflects nature, we can infer the following: during dry
236 conditions, runoff production does not begin significantly until the evaporation efficiency is
237 about 0.3, and the runoff ratio is not especially sensitive to changes in soil moisture (i.e., relative
238 to the sensitivity of evaporation efficiency to soil moisture) until the soil is wet enough to
239 support an evaporation efficiency of about 0.6, at which point the runoff ratio becomes strongly
240 sensitive to soil moisture. Evaporation efficiency, in contrast, is strongly sensitive to soil
241 moisture variations during dry conditions and is much less sensitive to soil moisture variations
242 during wet conditions.

243

ii. Bearing of curve position on simulated streamflows. Figure 6 is provided to illustrate how different curves in efficiency space translate to different simulations of streamflow. When the WBM is fitted with evaporation and runoff efficiency functions corresponding to the black curve in Figure 5b, the resulting simulation of streamflow is, as expected, quite accurate (Figure 6a). Again, though, any curve lying in the dark red region of the plot will produce a reasonably accurate simulation. The right panel of Figure 6b, for example, shows the WBM streamflows associated with the white curve in the left panel of Figure 6b, a curve which lies within the dark red region. The simulated streamflows are only slightly less accurate than those in Figure 6a, and given presumed uncertainties in the streamflow observations and in the meteorological forcing data used, one simulation cannot be said to be truly superior to the other.

Worth mentioning here is that the skill levels shown in Figure 5b are robust. If only the first half of the meteorological forcing and streamflow observations are used to establish the skill scores, the resulting optimal curves (not shown) are found to be very similar to those in the figure and result in similarly accurate streamflow simulations during the second half of the study period.

Figure 6c shows that when the efficiency space curve passes through the northwest quadrant, the resulting streamflow simulations are grossly overestimated. It seems safe to conclude that this curve does not come close to representing nature. On the other hand, when the efficiency space curve lies far below the black curve, as in Figure 6d, the simulated streamflows are significantly underestimated. Figure 6e shows the results obtained for an insensitivity of runoff efficiency to soil moisture. The simulated streamflows do vary slightly from year to year, but as expected, their interannual variability appears significantly underestimated relative to observations.

267

268 e. Water Holding Capacity

269 The two main user-defined elements of the WBM are the evaporation efficiency and
270 runoff efficiency functions, which, as shown above, can be described together as a single curve
271 in efficiency space. The remaining user-prescribed WBM parameter is the water holding
272 capacity, C_w. In the simulations discussed above, C_w is set to 169 mm, based on a chosen depth
273 of 0.5 m, an assumed wilting point at a degree of saturation of 0.25, and a porosity of 0.45. For
274 completeness, it is worth examining how the WBM results vary with C_w.

275 The efficiency space skill analysis in Figure 5b is repeated in Figure 7 for depths of 0.1
276 m, 0.5 m, 1. m, and 1.667 m; again assuming a porosity of 0.45 and a wilting point of 0.25, these
277 depths correspond to C_w values of 34 mm, 169 mm, 338 mm, and 563 mm. Here the skill is
278 evaluated over all of the shaded basins in Figure 7a collectively, with the skill contribution from
279 each basin weighted by basin area.

280 Two features from Figures 7b-7e stand out. First, while the optimal (black) curves differ
281 somewhat between the four panels, the regions of high skill within efficiency space (the darker
282 red areas) are basically the same – the highest scoring curves tend to hug the x-axis for lower
283 values of $\lambda E/R_{net}$ and then climb sharply into the plot's northeast quadrant for higher values of
284 $\lambda E/R_{net}$. This basic curve shape is also seen for individual basins and other depths (not shown).
285 The estimation of a valid relationship between evaporation efficiency and runoff efficiency
286 functions is largely insensitive to the choice of C_w.

287 The second interesting feature of Figures 7b-7e is the fact that higher skill levels are
288 possible with a depth of 0.5 m or 1.0 m than with a depth of 0.1 m or 1.667 m. (The four panels

use the same color bar to illustrate skill, with levels that slightly differ from those used in Figure 5.) This result could be interpreted in terms of the depth of soil "actively" participating in the land-surface hydrological cycle – perhaps, in a gross sense, changes in the soil water storage within the top 0.5 m or 1 m of soil are most relevant to the interannual variability of surface hydrological fluxes. Curiously, this depth range is roughly consistent with nominal estimates of vegetation rooting depth. Such agreement is perhaps not a coincidence.

f. Discussion

The patterns shown in the efficiency space skill diagrams are entirely determined by the imposed observations-based precipitation and net radiation forcing and by the streamflow measurements used to evaluate the WBM output. Thus, the precipitation, net radiation, and streamflow observations, when examined jointly with the WBM, can be said to contain important and otherwise hidden information: (i) the fact that (based on the shape of the high scoring curves) evaporation's sensitivity to soil moisture variation is greater than that of runoff in dry situations, whereas the reverse is true in wet situations; and (ii) the soil depth of greatest relevance to the generation of hydrological fluxes is perhaps of the order of 0.5 m to 1 m (Figure 7). Most importantly, the high-scoring efficiency space curves suggest how evaporation and runoff efficiencies vary with each other in nature. This joint variation is critical to defining a basin's hydrological behavior, at least in terms of how much runoff it produces each year; any other version of the joint variation (i.e., any curve in efficiency space lying outside the high-scoring region) produces inaccurate streamflows.

The high scoring curves for individual basins (not shown) are similar to first order, and yet they do nonetheless show some differences – there is, of course, some spatial heterogeneity in how nature behaves. Preliminary analyses of the WBM results indicate, for example, that the efficiency space curves for mountainous basins tend to lie a little above those for flatter basins in the dry regime, supporting the intuitive notion that mountainous areas produce runoff more easily than flat areas. Also, preliminary analyses suggest that flatter areas can produce higher skill scores with deeper soil depths, consistent with the fact that deeper depths-to-bedrock are found in flatter areas (e.g., as seen in STATSGO data [NRCS, 2012]). While a proper analysis of the spatial heterogeneity of basin behavior is beyond the scope of this paper, the potential for such heterogeneity must be kept in mind when interpreting the results presented herein.

3. The Budyko Perspective

Budyko (1974), building on analyses by Schreiber (1904) and Ol'dekop (1911), pioneered the joint analysis of evaporation and runoff response to precipitation and net radiation forcing. Although Budyko's analyses are focused on climatic means – how, for example, annual mean evaporation varies with annual mean precipitation and net radiation – rather than on the short time scale (daily time steps or shorter) of relevance here, a comparison of Budyko's results with those seen in the efficiency space diagrams is of interest.

Budyko's equation for evaporation is:

$$E / P = [(R_{net}/P\lambda) \tanh (P\lambda/R_{net}) (1 - \cosh(R_{net}/P\lambda) + \sinh (R_{net}/P\lambda)]^{1/2} . \qquad (5)$$

330 This function is plotted in Figure 8a. If we let the variable D represent the dryness index (i.e., D
331 = R_{net} / λP), then Budyko's equation can be rewritten:

332 $$\lambda E / R_{net} = f(D); \quad (6)$$

333 that is, the climatological evaporation efficiency is some complex function of dryness index.
334 Now consider that for long-term averages, soil moisture variations in the water budget equation
335 are irrelevant, so that the water balance can be written:

336 $$P = Q + E. \quad (7)$$

337 Combining (6) and (7) then provides an equation for the climatological runoff efficiency:

338 $$Q / P = 1 - D\, f(D). \quad (8)$$

339 Together, (6) and (8) allow the equivalent of Budyko's climatological equation to be
340 plotted in efficiency space. The idea is simple: a given value of D leads, through the two
341 equations, to a unique pairing of $\lambda E/R_{net}$ and Q/P and thus to a unique point in efficiency space.
342 Examining a wide range of D values leads to a wide range of points, allowing the full curve to be
343 plotted within the space.

344 This curve is shown in white in Figure 8b. The color-coded skill levels appearing in the
345 plot are taken from Figure 7c (i.e., they represent the WBM-based skill levels over the gray-
346 colored basins in Figure 7a for a depth of 0.5 m). Two aspects of the Budyko curve in Figure 8b
347 stand out. First, the curve has the same basic shape as the WBM-based high-scoring curves;
348 indeed, Budyko's curve falls within the high-scoring region of the space, though only just barely.
349 Thus, even though Budyko's equation is built to capture climatological efficiencies and
350 implicitly (if empirically) accounts, for example, for the effects of seasonal phase differences in

351 the precipitation and net radiation forcing (Koster et al. 2006), the equation produces a curve in
352 efficiency space that is not grossly inconsistent with the WBM-based curves, which focus on the
353 short time-scale problem.

354 Second, the Budyko-based curve is nevertheless seen to overestimate evaporation
355 efficiency at the wet end, if we assume that either the optimal curve (in black) or the median
356 curve within the high scoring region of the space is most correct. When the soil is fully wet (at
357 the northeast corner of the plot), the Budyko curve implies that all of the net radiation will be
358 converted into latent heat flux, i.e., that the land surface provides no resistance whatsoever to
359 evaporation, even given the complexity of transpiration pathways through the vegetation. If such
360 an extreme were true for climatological means, arguably it would also need to be true at any
361 smaller time step, thus contradicting the median behavior of the WBM results. Some past
362 studies (e.g., Koster et al. 2006) have indeed found observational evidence that Budyko's
363 estimates of climatological evaporation at the wet end are too high.

364

365 **4. Relevance to Land Model Development**

366

367 a. LSM Evaluation and the Model Dependence of Soil Moisture

368 When evaluating the realism of an LSM, a key element to consider is the manner in
369 which it allows variations of soil moisture to affect the magnitudes of evaporation and runoff
370 fluxes. Figure 9 shows the diagnosed relationship between evaporation efficiency and soil
371 moisture for three separate and representative LSMs (names withheld). The abscissa of each

point in each scatter plot corresponds to a monthly-averaged root zone soil moisture, and the ordinate of each point is the ratio of monthly-averaged evaporation to monthly-averaged net radiation; the data, taken from archived output of multidecadal offline simulations, are plotted here for points in and about the CONUS area during warm season months (May – September). A significant amount of scatter is seen in these plots, reflecting the fact (Section 1) that LSMs compute evaporation and runoff using a complex set of interacting parameterizations and thus do not generally use such efficiency functions directly. (These plots indeed demonstrate that the WBM is substantially simpler than a standard LSM.) Nevertheless, despite the scatter, a first order relationship between $\lambda E/R_{net}$ and soil moisture can be seen for each model.

Figure 9 shows why evaluating such relationships directly can be difficult. Within Model A, for example, the minimum soil moisture achieved varies spatially due to spatial variations in soil and vegetation properties, such as wilting point; for this model, there is no single relevant evaporation efficiency function. More importantly, evaluating an LSM's efficiency function is difficult because the LSM's soil moisture variable is model-specific, more of an "index of wetness" than a quantity that can be compared directly to observations (Koster et al. 2009). Such model dependency is reflected, for example, in the contrasting soil moisture ranges of Models B and C. Soil moisture and evaporation observations at the large scale are notoriously sparse, but even if they were plentiful, using them to evaluate a model's evaporation efficiency relationship would be hampered substantially by this model dependency.

The efficiency space framework, however, provides an elegant means for avoiding this problem. According to this framework, the model-dependent shapes of an LSM's $\lambda E/R_{net}$-vs-W and Q/P-vs-W relationships are of secondary importance; the key relationship defining the LSM's hydrological behavior is that between evaporation efficiency and runoff efficiency, a

relationship that can be examined without the direct consideration of soil moisture. The LSM can thus be evaluated by comparing its effective efficiency space curve against the corresponding observations-based curve (e.g., the black curve in Figure 7c).

b. Using Efficiency Space to Evaluate State-of-the-Art Land Models

Using archived model output to determine directly the efficiency space curve that best characterizes a given LSM's behavior comes with difficulties. Chief among these is the lack of an optimal averaging period over which to accumulate the evaporation, runoff, precipitation, and net radiation diagnostics. If the averaging period is too short (e.g., a week or so), then the runoff associated with baseflow, with its associated time delay, is no longer tied strongly to the concurrent precipitation. If, on the other hand, the averaging period is too long (e.g., a season), seasonal variability in soil moisture will confound the curve's derivation – most of the runoff, for example, may be produced during a month of snowmelt, when soil moisture is high, whereas most of the evaporation may be produced a couple of months later, when the soil moisture is lower. Monthly diagnostics are a reasonable compromise, but even these, to some extent, are subject to such problems.

To compute the LSM-based efficiency space curves shown in Figure 10, a different approach is employed, one that avoids these problems and has the additional advantage of producing a simple, smooth curve in the space. In essence, the approach used to compute the highest scoring (black) curve in efficiency space in Figures 5b and 7 is repeated, this time using the full LSM's time series of streamflow values (rather than the observations) as the "truth" upon which to base the RMSE metric. The precipitation and net radiation forcing applied to the WBM

417 is derived from that used originally by the LSM in question, when it was generating its archived
418 data. Thus, whereas the black curves in Figures 5b and 7c can be said to be in line with how
419 nature works, the curve produced using the LSM-based "streamflow truth" can be said to be in
420 line with how the LSM works – it can be said to characterize, in a gross sense, the LSM's
421 underlying relationship between evaporation and runoff efficiencies.

422 The left panel of Figure 10a shows, in white, the optimal efficiency space curve
423 established for a specific LSM (name withheld) based on its streamflow generation in the gray-
424 colored basins of Figure 7a; plotted around it are the corresponding observations-based color-
425 coded skill curves from Figure 7c. The curve for the LSM clearly lies above the high-scoring
426 region of the space (where "nature's curve" lies), especially toward the dry end. Based on the
427 discussion of Figure 6c, this LSM can be expected to compute too much runoff. Sure enough,
428 the right two panels in Figure 10a show that the annual runoff ratios computed by the full LSM,
429 as determined from the archived diagnostics, are too large compared to observations, with the
430 largest errors seen toward the western half of CONUS.

431 Note that a traditional interpretation of the comparison in the rightmost panels of Figure
432 10a might be that the runoff formulation in this LSM is deficient. This may be true, but the
433 efficiency space framework allows the deficiency to be stated in a potentially more helpful way:
434 at the drier end, this LSM's runoff efficiency is too large relative to its evaporation efficiency.

435 Figure 10b shows the optimal curve obtained for a different LSM. Here the LSM-based
436 curve falls below the high-scoring region, and, consistent with the discussion of Figure 6d, the
437 resulting streamflows are underestimated relative to observations (rightmost panels). Again, the
438 efficiency space framework allows us to describe this LSM's deficiency in terms of its joint

evaporation and runoff behavior rather than solely in terms of its runoff formulation. Correcting this LSM – bringing the white curve into the high-scoring region of the plot – could be a matter of improving its runoff formulation (increasing Q/P for a given soil moisture), but alternatively, and perhaps just as effectively, it could also result from improvements in the evaporation formulation (decreasing $\lambda E/R_{net}$ for a given soil moisture).

Six other state-of-the-art LSMs were examined in the same way. The results (not shown) indicate that most of these LSMs are similar to Model Y in Figure 10b – most of the LSMs are characterized by efficiency space curves that fall below the high scoring region.

c. Potential for Simple Land Model Tuning

Model Y in Figure 10b is in fact the Catchment-CN LSM (Koster et al. 2014), a merger of the energy and water balance framework of the Catchment LSM (Koster et al. 2000) and the dynamic phenology (prognostic biogeochemistry) components of the NCAR/DOE CLM4 dynamic vegetation model (Oleson et al. 2010). As suggested by the figure, this relatively new LSM could benefit from further development or tuning of its evaporation and runoff formulations.

When faced with the discrepancy in Figure 10b (i.e., the LSM's optimal curve, in white, lying outside the plot's high scoring region), one could envision two basic model development approaches. The first approach is simply to keep improving the evaporation and runoff formulations until the position of the white curve moves upward into the high scoring region. In principle this is the desirable approach, for if it is successful, the LSM's curve would end up being in the right place for the right reasons. With this approach, knowledge of the relative

461 strengths of the evaporation and runoff formulations – knowing which formulation can be trusted
462 more, based on the breadth of analysis underlying each – would point to the formulation that
463 currently acts as the "weak link" in the simulation of hydrological fluxes and is thus the most
464 appropriate target for focused development. The position of the white curve relative to the high
465 scoring region would indicate a direction for the change in the model formulation.

466 Improving a land model formulation through first principles, however, is far from trivial;
467 depending on the accuracies desired, it may encompass an entire scientific career. Given time
468 constraints and other considerations, the second model development approach – tuning an LSM
469 formulation using the observations-based efficiency space curves – may, at least for some
470 applications, be advantageous.

471 The tuning approach in its simplest form is illustrated in Figure 11. The first step
472 involves deciding whether the LSM's evaporation or runoff formulation can be trusted the most.
473 Compared to some LSMs, the Catchment-CN LSM has a relatively complex runoff calculation;
474 still, the complexity built into the runoff formulation pales in comparison to that built into the
475 evaporation formulation, and thus the latter is considered, for purposes of this demonstration, to
476 be more trustworthy. Plotted in Figure 11a are monthly (May through September) values of
477 $\lambda E/R_{net}$ versus monthly values of root zone soil moisture. The scatter here is similar to that
478 shown in Figure 9; again, though, a first-order relationship can be seen between the two
479 quantities. The red curve in Figure 11a, obtained by averaging the $\lambda E/R_{net}$ values over soil
480 moisture bins, is used in this exercise to characterize the relationship.

481 Figure 11b shows the observations-based efficiency space curves from Figure 7c. The
482 black dotted curve is taken to be the "target" efficiency space curve for the tuned land model;

this curve, which falls wholly within the high scoring region, is used instead of the solid black curve because it accounts for a larger range of $\lambda E/R_{net}$ values and thus works better for the tuning procedure. With the curves in Figures 11a and 11b in place – the first showing a relationship between $\lambda E/R_{net}$ and W, and the second showing one between $\lambda E/R_{net}$ and Q/P – a relationship between Q/P and W is easily derived (Figure 11c).

For the next step in the demonstration, the somewhat complex runoff formulation in the Catchment-CN LSM is completely stripped out of the model and replaced by a much simpler formulation, one in which the total runoff for a given time step is computed by multiplying the time step's incident precipitation by a Q/P ratio taken from Figure 11c, based on the current value of root zone soil moisture. The results of this tuning exercise are shown in Figure 12. The tuning, in general, leads to improved streamflow simulation. The tuned LSM (Figure 12b) – which still, by the way, maintains most of its complexity (its diurnally-varying energy balance calculations, its spatial variation in vegetation, and so on) – produces annual runoff ratios that are, for the most part, more in line with the observations (Figure 12c) than those obtained with the original LSM (Figure 12a).

This is, of course, a simple and rather heavy-handed sample demonstration. A more refined tuning exercise, one that isolates, for example, the particular facets of the runoff formulation that are known to be especially weak or arbitrary and then modifies only those facets, would in fact be much more satisfying and justifiable. Also, tuning the model separately for each basin would presumably produce better basin-by-basin results than those shown in Figure 12b. Given the difficulty of improving LSMs through first principles, a comprehensive tuning study may lead to a straightforward and more reasonable way of inducing an LSM to produce a more realistic simulation of hydrology. The present example serves merely to

illustrate how the information content of the observations-based efficiency space curves (e.g., Figure 7c) can work its way into an LSM formulation.

5. Summary and Discussion

The typical land surface model, or LSM, is constructed as follows: the modeler produces the best evaporation formulation and the best runoff formulation possible and then combines them (along with formulations for various other land processes) into a single system. The modeler then tests this LSM against observations of (for example) surface turbulent fluxes and local streamflow rates and uses the test results, in conjunction with his or her knowledge of current model weaknesses, to adjust the individual LSM components. While on the surface this seems reasonable, the evaporation and runoff processes within an LSM are often developed and examined independently, thereby ignoring the fundamental truth emphasized in this paper: a first order understanding of the hydrological behavior simulated by an LSM requires the *joint* analysis of its evaporation and runoff formulations, so that the variations of evaporation and runoff efficiencies with respect to each other are, to first order, understood.

To explore this joint variation, and to provide a potentially useful tool for developing, evaluating, and improving LSM formulations, the present paper introduces the concept of efficiency space. A simple water balance model (WBM) used in conjunction with precipitation, net radiation, and streamflow observations produces curves in efficiency space that illustrate, in a concise and simple way, how evaporation and runoff efficiencies tend to vary with each other in nature. Considering together all of the gray basins in Figure 7a, nature's curve in efficiency space lies somewhere within the dark red region of one of the lower panels of Figure 7. In its

528 own way, such a curve is conceptually powerful, as it captures, to first order, the net effect of
529 land surface processes on annual hydrological fluxes.

530 Of course, such an analysis strategy does not come without its caveats. As discussed in
531 the text, for example, the WBM's structure is necessarily very simple, being designed to capture
532 only the most important controls on the surface moisture fluxes. Another caveat, not yet
533 discussed, involves the "control volume" assumed for the analysis. One holistic view of
534 evaporation from the land surface interprets it as the net result of a variety of interacting
535 processes in the soil, in the vegetation canopy, and in the atmospheric boundary layer, all
536 evolving together (Betts 2004). Betts (2004) argues that addressing the evaporation problem
537 through only a subset of these processes can have significant limitations. While this is certainly
538 a valid and useful view, the analysis of evaporation in the present paper effectively centers on an
539 alternative control volume, one that represents the soil and vegetation from the top of the canopy
540 to some depth below the surface (say, a meter or two). In a sense, this control volume represents
541 the operating world of the land surface model component of a climate model – a system that is
542 forced from without by precipitation and incoming radiation and that must partition this
543 incoming water and energy into a variety of fluxes and storage changes. The choice here of this
544 alternative control volume, with its neglect of a detailed consideration of boundary layer and
545 cloud processes, is guided – and is in fact made necessary – by the particular focus of this paper,
546 namely, the examination of the dual role of soil moisture in controlling both evaporation and
547 runoff (streamflow) production.

548 Keeping these caveats in mind, the curve representing nature's behavior in Figure 7c can
549 be used as a target for LSM development. By noting where a model's efficiency curve lies
550 relative to that of nature, the modeler can adjust either the evaporation or the runoff formulation

551 – presumably whichever one is known to be weaker – in the right direction to bring the curves in
552 line. The result should be an overall improvement in the simulation of streamflow and
553 evaporation. Using the efficiency curve as a target allows the model developer to sidestep a
554 commonly encountered problem in the evaluation of evaporation and runoff formulations: the
555 fact that soil moisture in an LSM is more of a model-dependent index of wetness than a quantity
556 that can be compared directly to observations.

557 While focusing on efficiency space takes soil moisture out of the LSM development
558 problem in a clean way, this must not be construed to imply that measurements of soil moisture
559 are not useful for hydrological analysis and LSM development. The advent of satellite-based
560 soil moisture measurements, particularly L-band estimates from the SMOS (Kerr et al. 2010) and
561 SMAP (Entekhabi et al. 2010) sensors, opens the door to new and potentially powerful research
562 into hydrological processes, including that associated with joint evaporation and runoff
563 efficiencies. Consider, for example, the evaporation efficiency and runoff efficiency curves in
564 Figure 13. The shapes of the functions here are arbitrarily drawn, but together they capture the
565 basic shape of nature's curve in efficiency space. As suggested in the figure, evaporation-
566 induced changes in soil moisture during non-rainy periods should be smaller and more variable
567 in time when the soil is in the drier regime, whereas rainfall-induced changes should be smaller
568 and more variable in time in the wetter regime. One can imagine that by sorting soil moisture
569 changes in the new satellite data into subsets associated with rainy and non-rainy periods and
570 then examining these changes in the context of efficiency function pairings that correspond to
571 nature's efficiency space curve, one could establish where on that curve nature is operating at a
572 given time – a non-traditional and potentially very useful way to characterize the current "overall
573 hydrological state" of the measured region. Perhaps the analysis of the remotely sensed data

could even constrain better the curve's location. Such research, of course, would have to address a number of issues, including the limited vertical penetration (relative to the root zone) of the L-band signal; still, the possibilities are intriguing.

In closing, it is worth noting again that the joint consideration here of evaporation and runoff processes is in some ways not at all new, as it echoes, for example, the landmark work of Budyko (1971) and his predecessors. These earlier works focused on climatological means; the present paper, in a sense, discusses the ideas in the context of the short time scale and thus in the context of the operation of LSMs. The hope here is that the efficiency space framework, which succinctly captures the evaporation-runoff connection, can be used to promote LSM development. In the long run, such development should lead to improvements in overall Earth system model performance.

Acknowledgments. The author is deeply grateful to Sarith Mahanama and Greg Walker for their contributions to this work. Conversations with Max Suarez, Chris Milly, Guido Salvucci, and Siegfried Schubert about the ideas presented herein are also greatly appreciated. This work was supported by the Earth System Science Program of the NOAA/Climate Program Office and by the NASA Terrestrial Hydrology Program.

References

Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier, 2005: Twentieth-century drought in the conterminous United States. J. Hydrometeor., 6, 985-1001.

596 Best, M. J., and Co-authors, 2014: The PALS Land Surface Model Benchmarking Evaluation
597 Project (PLUMBER). Presentation at the AMS annual meeting, Atlanta, Georgia, USA,
598 February 2014.

599 Betts, A. K., 2004: Understanding hydrometeorology using global models. *Bull. Amer. Met.*
600 *Soc*., **85**, 1673-1688.

601 Boone, A., F. Habets, J. Noilhan, and coauthors 2004: The Rhone-Aggregation Land Surface
602 Scheme Intercomparison Project: An Overview, *J. Climate*, **17**, 187-208.

603 Bowling L. C., D. P. Lettenmaier, B. Nijssen and coauthors 2003: Simulation of high latitude
604 hydrological processes in the Torne-Kalix basin: PILPS Phase 2e 1: Experiment
605 description and summary intercomparison, *Global and Planetary Change*, **38**, 1-30.

606 Budyko, M. I., 1974: *Climate and Life*. Academic Press, New York, 508 pp.

607 Chen, T. H., and Coauthors, 1997: Cabauw experimental results from the Project for
608 Intercomparison of Land-surface Parameterization Schemes (PILPS). *J. Climate,* **10,**
609 1194–1215.

610 Delworth, T.L., and S. Manabe, 1989: The influence of soil wetness on near-surface
611 atmospheric variability. *J. Clim*., **2**, 1447-1462.

612 Dirmeyer, P. A., A. J. Dolman, and N. Sato, 1999: The pilot phase of the Global Soil Wetness
613 Project. *Bull. Amer. Meteor. Soc*., **80**, 851-878.

614 Dirmeyer, P. A., 2000: Using a global soil wetness dataset to improve seasonal climate
615 simulation. *J. Climate*, **13**, 2900-2922.

616 Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: GSWP-2
617 Multimodel Analysis and Implications for Our Perception of the Land Surface, *Bul.*
618 *Amer. Meteo. Soc.*, **87**(10),1387-1397 DOI: 10.1175/BAMS-87-10-1381

619 Entekhabi, D., and Co-authors, 2010: The Soil Moisture Active Passive (SMAP) mission. Proc.
620 IEEE, 98, 704-716.

621 Gupta, S. K., Stackhouse, P.W., Jr., S.J. Cox, J. C. Mikovitz, T. Zhang, 2006: 22-Year Surface
622 Radiation Budget Data Set. GEWEX News, V 16, No. 4, November, 12-13.

623 Henderson-Sellers, A., Z.-L. Yang, and R. E. Dickinson, 1993: The Project for Intercomparison
624 of Land-surface Parameterization Schemes. *Bull. Amer. Meteor. Soc.,* **74,** 1335–1349.

625 Kerr, Y. H., and Co-authors, 2010: The SMOS mission: New tool for monitoring key elements of
626 the global water cycle. Proc. IEEE, 98, doi:10.1109/JPROC.2010.2043032.

627 Koster, R. D., and P. C. D. Milly, 1997: The interplay between transpiration and runoff
628 formulations in land surface schemes used with atmospheric models. *J. Climate*, **10**,
629 1578-1591.

630 Koster, R. D., M. J. Suarez, and M. Heiser, 2000: Variance and predictability of precipitation at
631 seasonal-to-interannual timescales. *J. Hydrometeor.,* **1,** 26–46.

632 Koster, R. D., B. M. Fekete, G. J. Huffman, and P. W. Stackhouse, 2006: Revisiting a
633 hydrological analysis framework with International Satellite Land Surface Climatology
634 Project Initiative 2 rainfall, net radiation, and runoff fields. J. Geophys. Res., 111,
635 D22S05, doi:10.1029/2006JD007182.

636 Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based
637 approach to modeling land surface processes in a general circulation model: 1. Model
638 structure, *J. Geophys. Res.*, **105**(20), 24,809– 24,822.

639 Koster, R. D., Z. Guo, R. Yang, P. A. Dirmeyer, K. Mitchell, and M. J. Puma, 2009b: On the
640 nature of soil moisture in land surface models. *J. Climate*, *22*, 4322-4335.

641 Koster, R. D., and S. P. P. Mahanama, 2012: Land surface controls on hydroclimatic means and
642 variability. J. Hydromet., 13, 1604-1620.

643 Koster, R. D., G. K. Walker, G. J. Collatz, and P. E. Thornton, 2014: Hydroclimatic controls on
644 the means and variability of vegetation phenology and carbon uptake. Submitted to *J.*
645 *Climate*.

646 Mahanama, S., B. Livneh, R. Koster, D. Lettenmaier, and R. Reichle, 2012: Soil moisture, snow,
647 and seasonal streamflow forecasts in the United States. J. Hydromet., 13, 189-203.

648 Manabe, S., 1969: Climate and the ocean circulation, I, The atmospheric circulation and the
649 hydrology of the Earth's surface. *Mon. Wea. Rev.*, **97**, 739-774.

650 NRCS Soil Survey Staff, USDA (2012). General Soil Map (STATSGO2) [United States].
651 Available online at http://websoilsurvey.nrcs.usda.gov/.

652 Ol'dekop, E. M., 1911: Ob Isparenii s Poverkhnosti Rechnykh Basseinov (On evaporation from
653 the surface of river basins). Tr. Meteorol. Observ. Iur'evskogo Univ. Tartu, 4.

654 Oleson, K. W., and Co-authors, 2010: Technical description of version 4.0 of the Community
655 Land Model (CLM). NCAR Technical Note NCAR/TN-478+STR., National Center for
656 Atmospheric Research, P. O. Box 3000, Boulder, Colorado, 80307-3000.

657 Schreiber, P., 1904: Uber die Beziehungen zwishcen dem Niederschlag und der Wasserfuhrung
658 der Flusse in Mitteleuropa. *Z. Meteorol*., **21**, Pt. 10.

659 Seneviratne, S. I., D. Luthi, M. Litschi, and C. Schar, 2006: Land-atmosphere coupling and
660 climate change in Europe. *Nature*, **443**, 205-209.

661 Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A.
662 J. Teuling, 2010: Investigating soil moisture-climate interactions in a changes climate, A
663 review. *Earth-Science Reviews*, **99**, 125-161.

664 Shukla, J., and Y. Mintz, 1982: Influence of land-surface evapotranspiration on the earth's
665 climate. *Science,* **215,** 1498–1501.

666 Wood, E. F., and Co-authors, 1998: The Project for Intercomparison of Land-surface
667 Parameterization Schemes (PILPS) Phase 2(c): Red-Arkansas River Basin experiment, 1,
668 Experiment description and summary intercomparisons. *J. Global Planet. Change*, **19**,
669 115-135.

670 Yildiz, O. and A. P. Barros, 2007: Elucidating vegetation controls on the hydroclimatology of a
671 mid-latitude basin. J. Hydrol., 333, 431-448.

672 **List of Figures**

673

674 1. a. Illustration of how a representative evaporation efficiency function ($\lambda E/R_{net}$ vs. W, the red
675 curve on the right) and runoff efficiency function (Q/P vs. W, the blue curve on the right)
676 combine to produce a single curve in efficiency space (the green curve on the left). b. A
677 second sample illustration.

678 2. Schematic of the water balance model (WBM) used in this study. The forcing variables,
679 imposed daily, are precipitation (P) and net radiation (R_{net}); the WBM uses the imposed
680 evaporation and runoff relationships to compute the daily evaporation (E) and runoff (Q) as a
681 function of its prognostic water content (W). (Reproduced from Koster and Mahanama
682 [2012].)

683 3. a. Illustration of a typical WBM simulation. Evaporation and runoff efficiency curves (the red
684 and blue curves on the left, respectively) are implemented into the WBM, and the WBM is
685 driven with gridded observational data over a number of decades. In the course of the
686 simulation, the WBM produces gridded runoff values that are spatially aggregated to
687 hydrological basins for comparison with observed basin streamflows (right panel). b. A
688 second representative example.

689 4. Demonstration that different pairings of evaporation and runoff efficiency functions lead to
690 essentially equivalent streamflow simulations if their corresponding efficiency space curves
691 are the same. The three pairings of efficiency functions in the top row (evaporation efficiency
692 in red, runoff efficiency in blue) each map to the single efficiency space curve in the lower

693 left panel. WBM simulations with each pairing of functions correspondingly produce very
694 similar simulations of streamflow (lower right panel).

695 5. a. Three efficiency space curves, color-coded according to the degree to which corresponding
696 WBM simulations reproduce observed streamflows in the Upper Mississippi Basin. The red
697 curve corresponds to the most skillful simulation, and the blue curve to the least skillful
698 simulation. Skill is quantified with an RMSE metric, though numbers are not shown in order
699 to emphasize relative skill. b. Same, but with over 23000 color-coded curves overlain on the
700 plot. The solid black curve is the curve associated with the greatest simulation skill.

701 6. a. WBM simulation of streamflow in the Upper Mississippi (red curve in right panel)
702 associated with the highest-scoring curve in efficiency space (black curve in left panel; see
703 Figure 5b). The black curve in the right panel represents the observed streamflows for the
704 basin. b-e. Same, but for the WBM simulation associated with the white curve in the left
705 panel.

706 Figure 7. a. Basins (in gray) examined collectively in the analysis. See Mahanama et al. (2012)
707 for details on the basins and on the streamflow observations taken therein. b. Color-coded
708 efficiency space curves as in Figure 5b, but for an assumed 0.1 m soil depth and for the
709 collected set of basins. The black curve represents the curve associated with the highest
710 simulation skill. c. Same, but for an assumed 0.5 m soil depth. c. Same, but for an assumed
711 1m soil depth. d. Same, but for an assumed 1.667 m soil depth.

712 8. a. Budyko's semi-empirical relationship between the evaporation-precipitation ratio and the
713 dryness index, defined as $R_{net}/P\lambda$. b. Efficiency space curves from Figure 7c. A transform of
714 Budyko's relationship is shown in white.

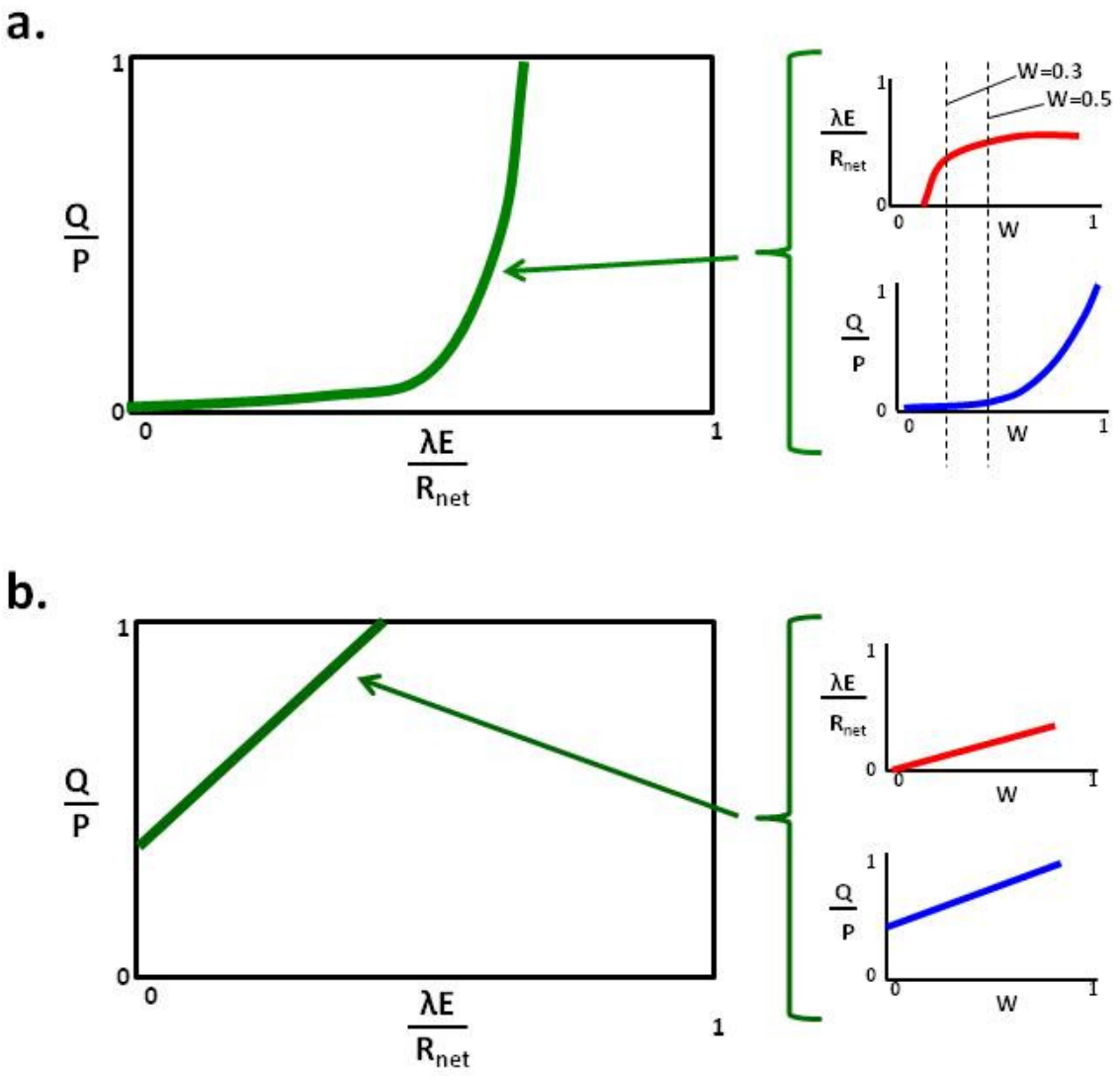
9. Scatter plots of evaporation efficiency (ratio of monthly-averaged E to monthly-averaged R_{net}) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by three state-of-the-art LSMs.

10. a. Efficiency space curve (in white) for a specific LSM ("Model X"), overlain on the color-coded curves from Figure 7c. To the right are plots of annual runoff ratio across several basins as computed from Model X diagnostics and from observations. b. Same, but for a different LSM ("Model Y").

11. a. Scatter plot of evaporation efficiency (ratio of monthly-averaged E to monthly-averaged R_{net}) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by the Catchment-CN LSM. b. Color-coded efficiency space curves from Figure 7c, with one of the curves highlighted as a heavy dotted line. c. Resulting relationship between runoff efficiency and root zone soil moisture.

12. a. Annual runoff efficiencies (ratio of annual basin-averaged runoff to annual basin-averaged precipitation) for several CONUS basins, as produced by the Catchment-CN LSM in a multidecadal offline simulation. b. Same, but for a tuned version of the Catchment-CN LSM (see text). c. Same, but for observations.

13. Sample evaporation efficiency and runoff efficiency functions, with highlighted regimes of interest.



734

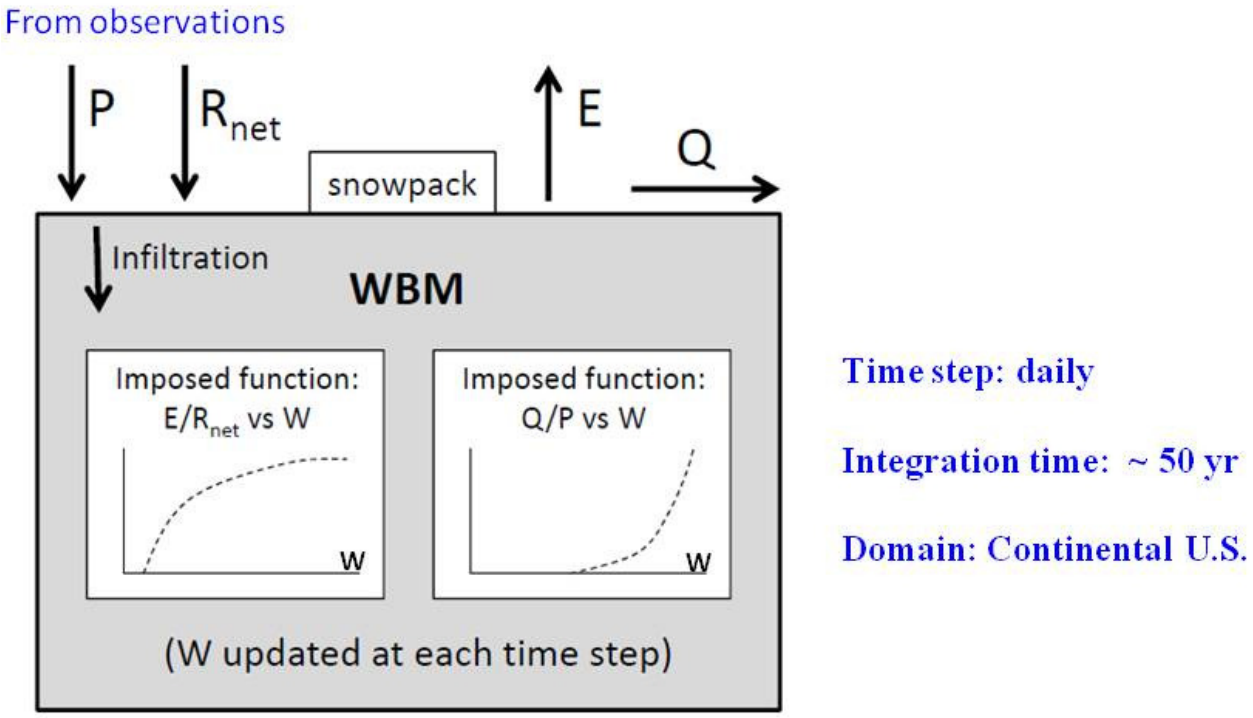
735

736 Figure 1. a. Illustration of how a representative evaporation efficiency function ($\lambda E/R_{net}$ vs. W,
737 the red curve on the right) and runoff efficiency function (Q/P vs. W, the blue curve on the right)
738 combine to produce a single curve in efficiency space (the green curve on the left). b. A second
739 sample illustration.

740

741

742



743

744

745

746

747 Figure 2. Schematic of the water balance model (WBM) used in this study. The forcing
748 variables, imposed daily, are precipitation (P) and net radiation (R_{net}); the WBM uses the
749 imposed evaporation and runoff relationships to compute the daily evaporation (E) and runoff
750 (Q) as a function of its prognostic water content (W). (Reproduced from Koster and Mahanama
751 [2012].)

752

753

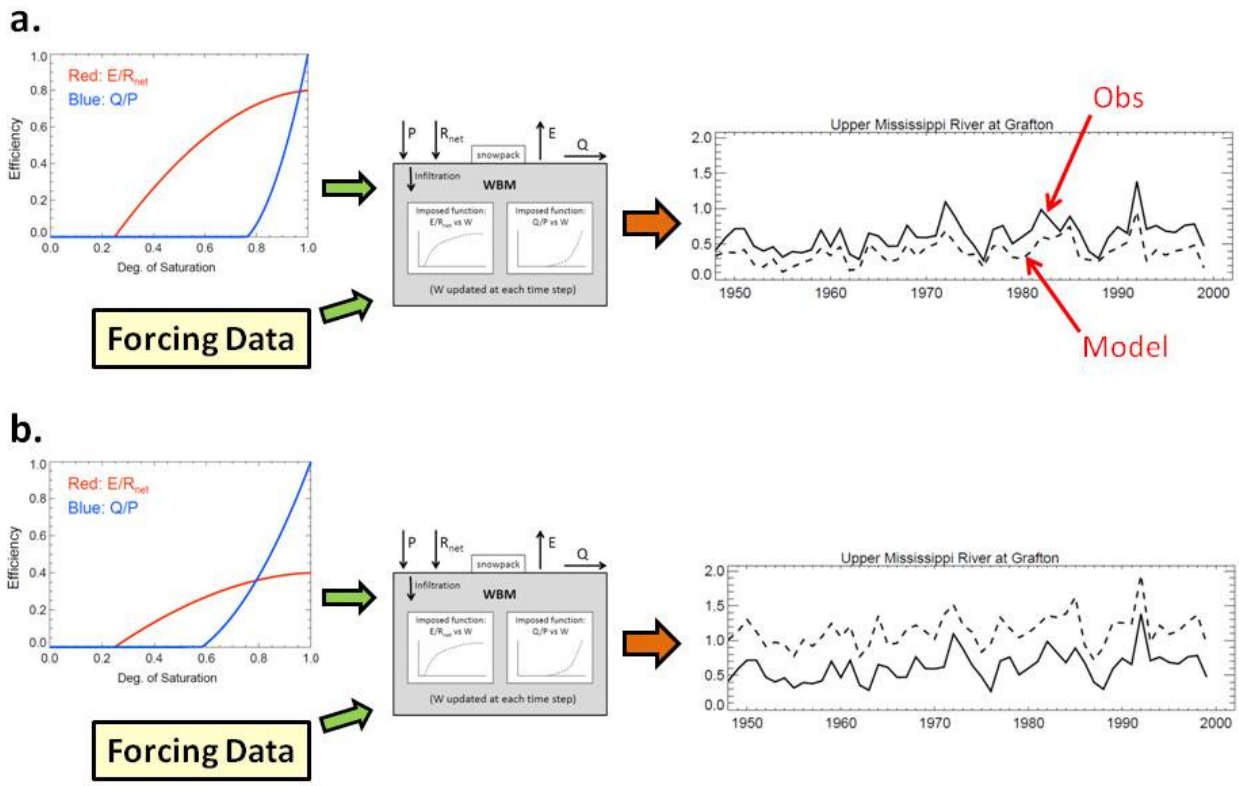


Figure 3. a. Illustration of a typical WBM simulation. Evaporation and runoff efficiency curves (the red and blue curves on the left, respectively) are implemented into the WBM, and the WBM is driven with gridded observational data over a number of decades. In the course of the simulation, the WBM produces gridded runoff values that are spatially aggregated to hydrological basins for comparison with observed basin streamflows (right panel). b. A second representative example.

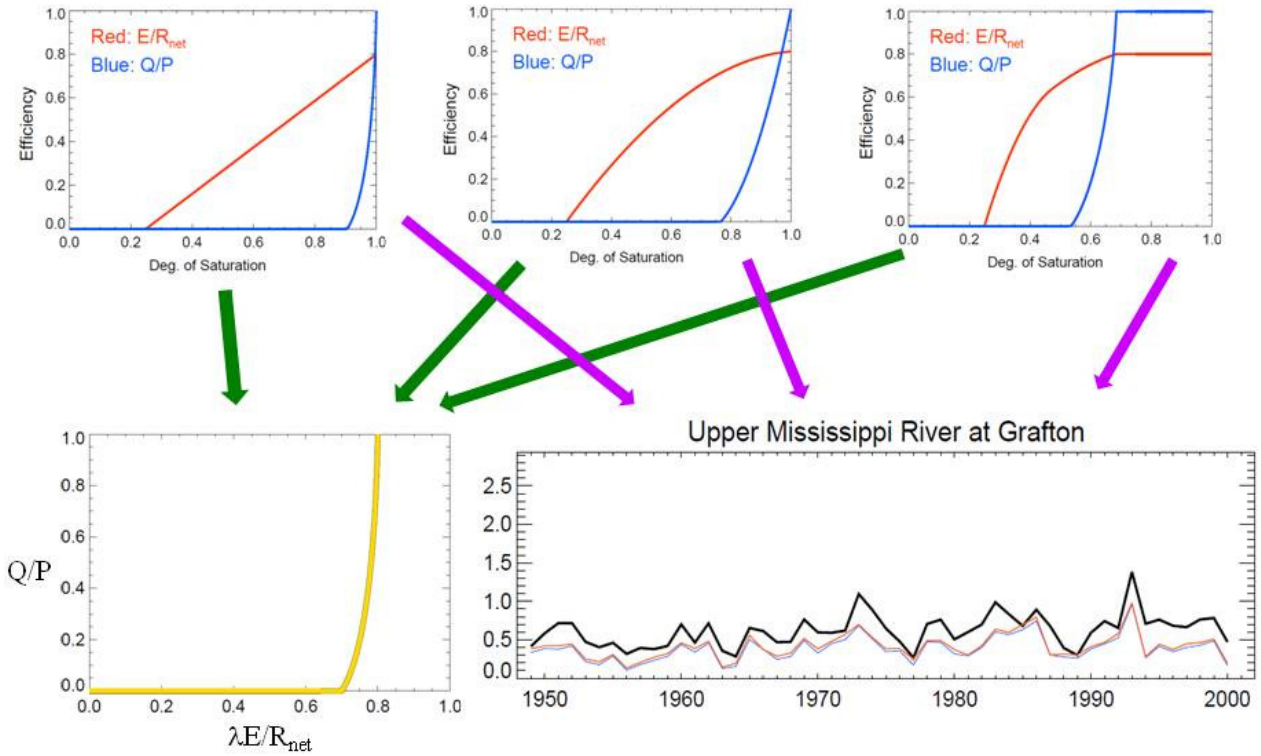


Figure 4. Demonstration that different pairings of evaporation and runoff efficiency functions lead to essentially equivalent streamflow simulations if their corresponding efficiency space curves are the same. The three pairings of efficiency functions in the top row (evaporation efficiency in red, runoff efficiency in blue) each map to the single efficiency space curve in the lower left panel. WBM simulations with each pairing of functions correspondingly produce very similar simulations of streamflow (lower right panel).

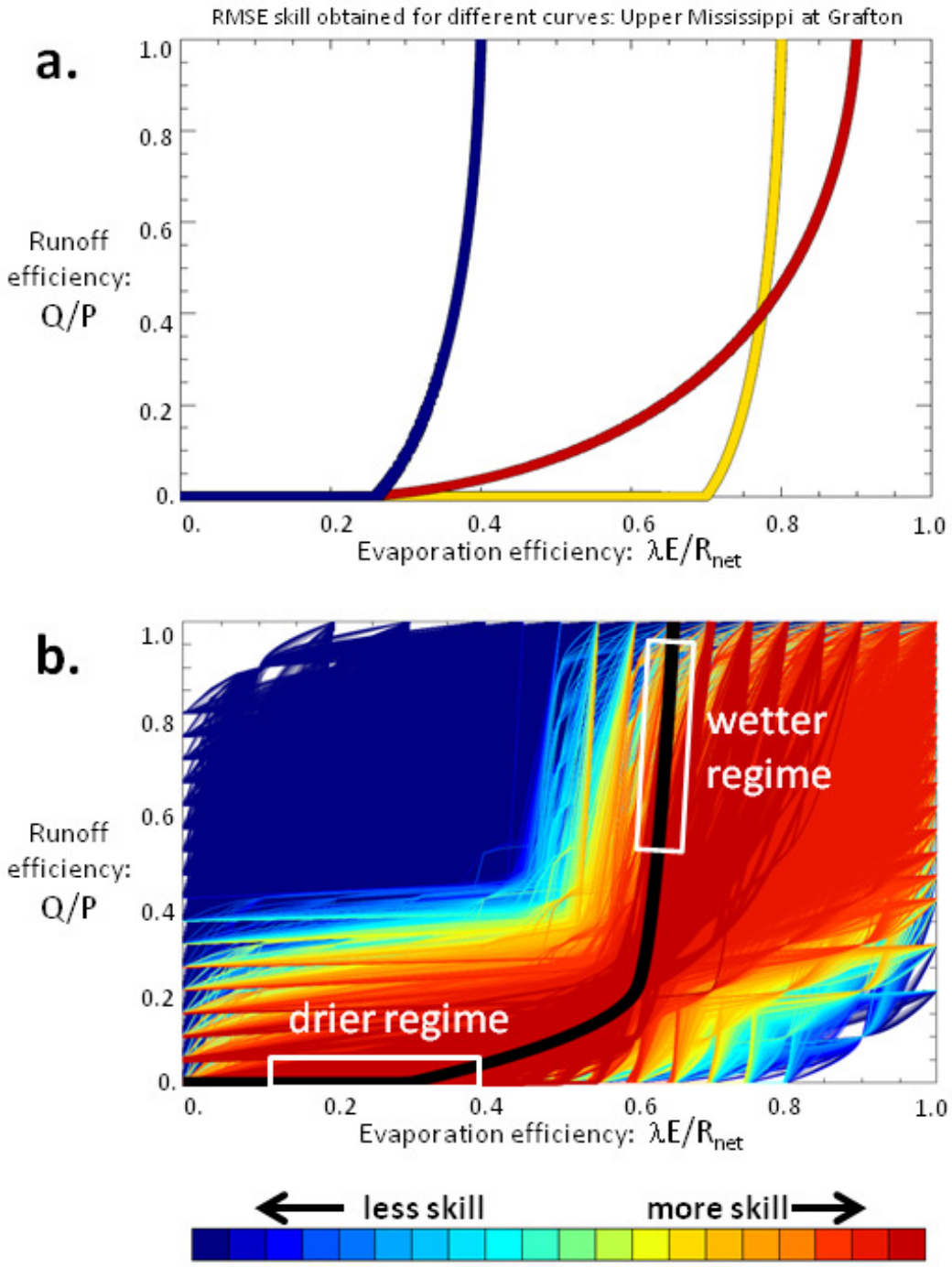


Figure 5. a. Three efficiency space curves, color-coded according to the degree to which corresponding WBM simulations reproduce observed streamflows in the Upper Mississippi Basin. The red curve corresponds to the most skillful simulation, and the blue curve to the least skillful simulation. Skill is quantified with an RMSE metric, though numbers are not shown in order to emphasize relative skill. b. Same, but with over 23000 color-coded curves overlain on the plot. The solid black curve is the curve associated with the greatest simulation skill.

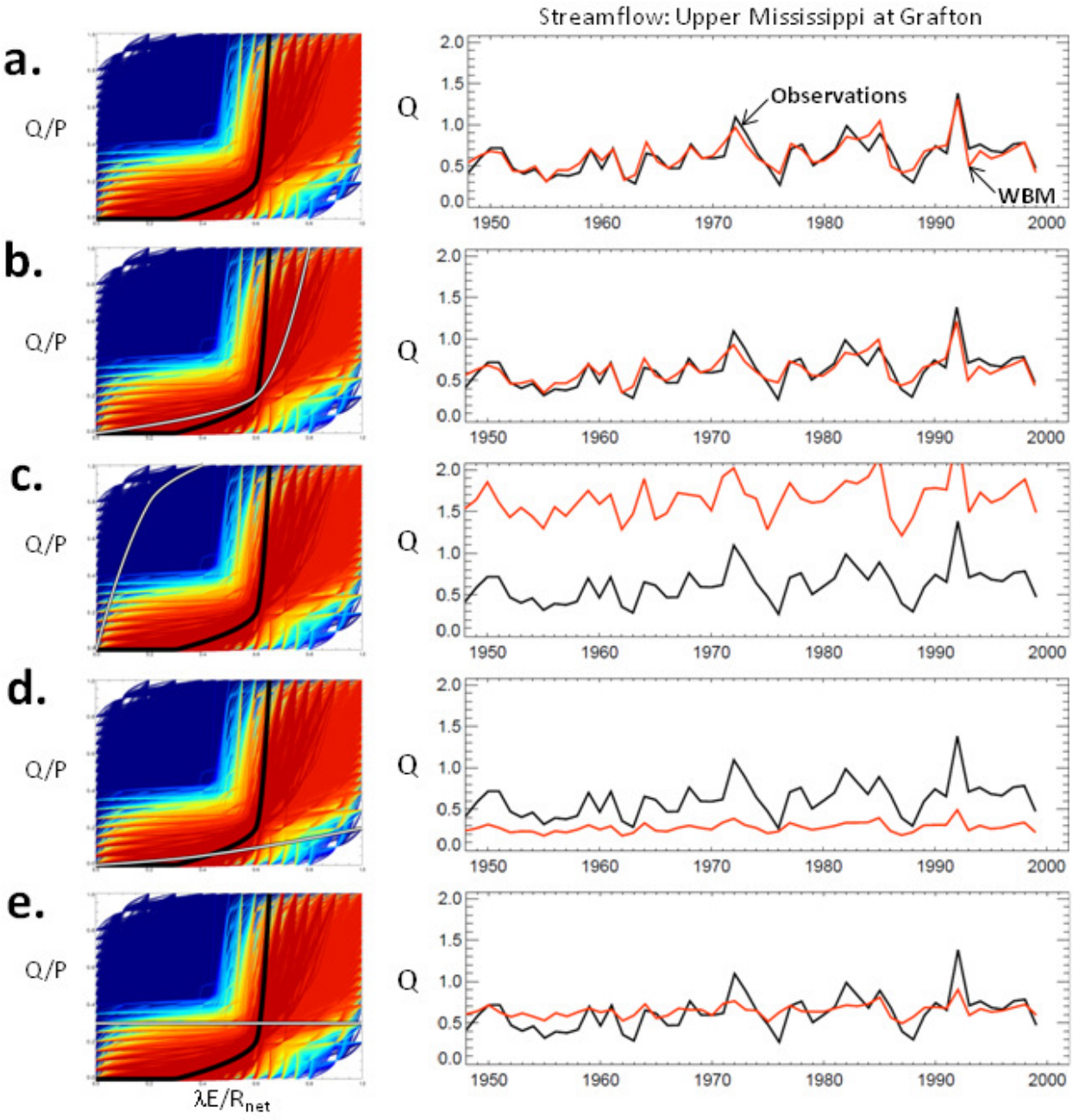


Figure 6. a. WBM simulation of streamflow in the Upper Mississippi (red curve in right panel) associated with the highest-scoring curve in efficiency space (black curve in left panel; see Figure 5b). The black curve in the right panel represents the observed streamflows for the basin. b-e. Same, but for the WBM simulation associated with the white curve in the left panel.

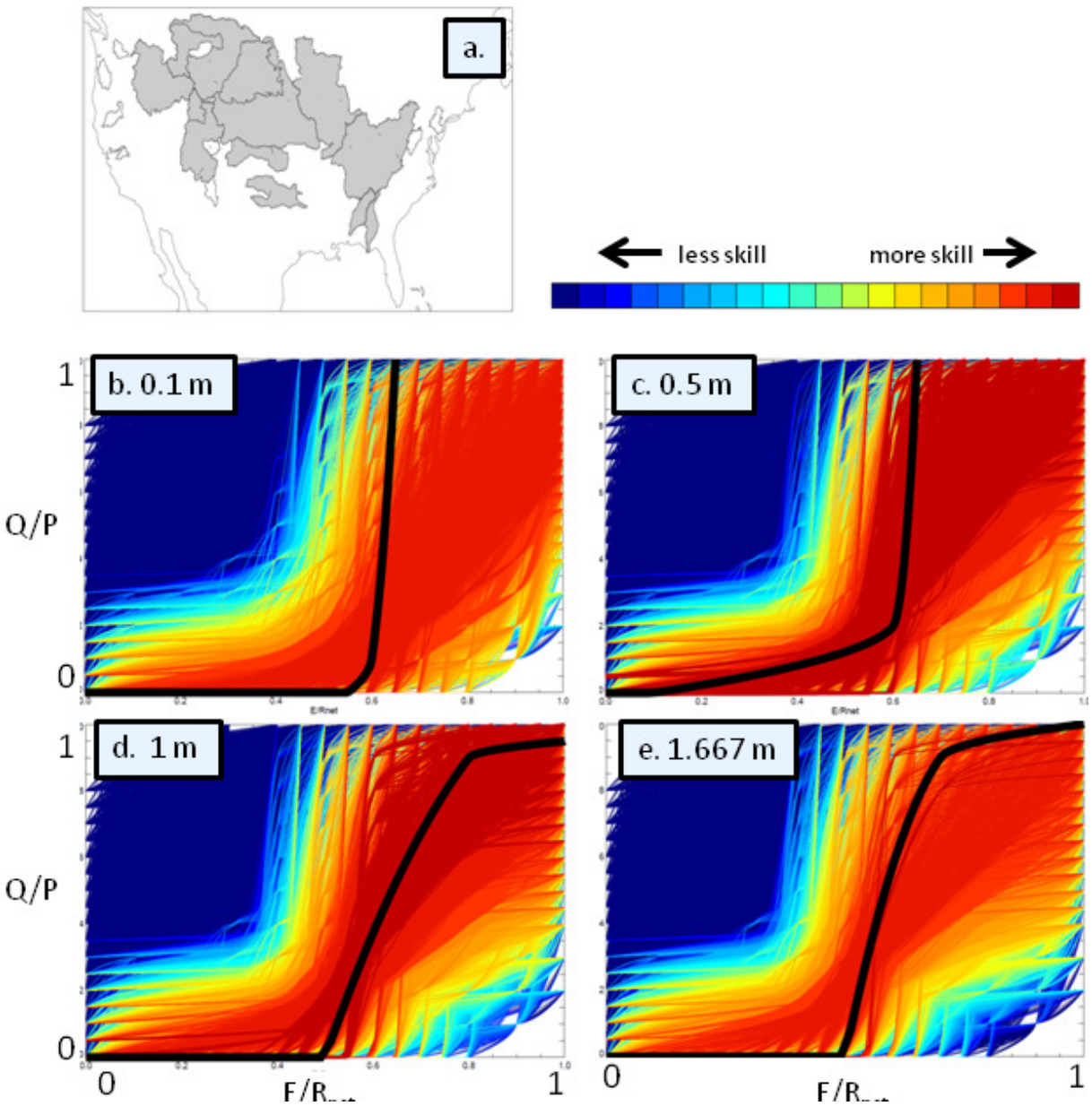
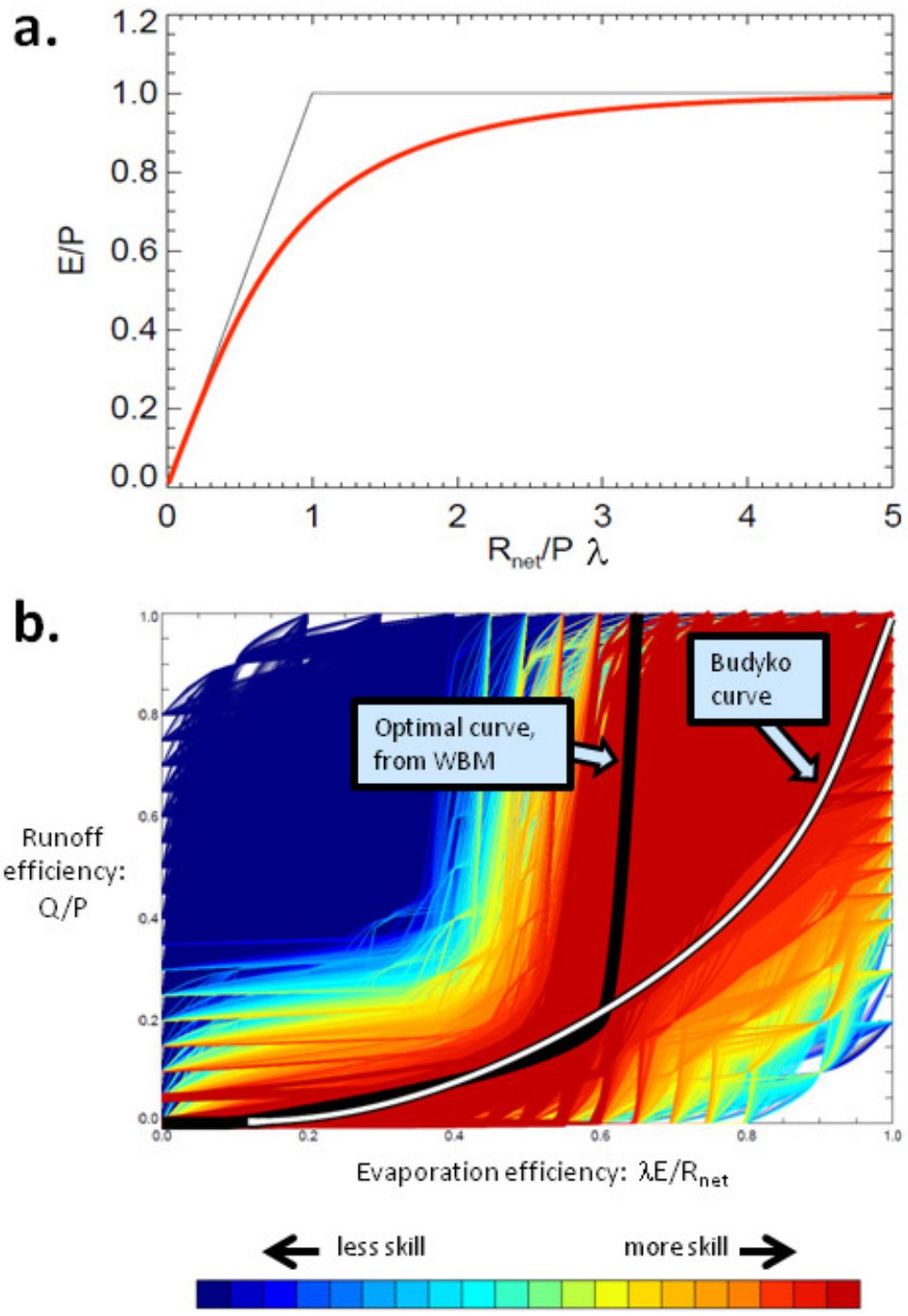


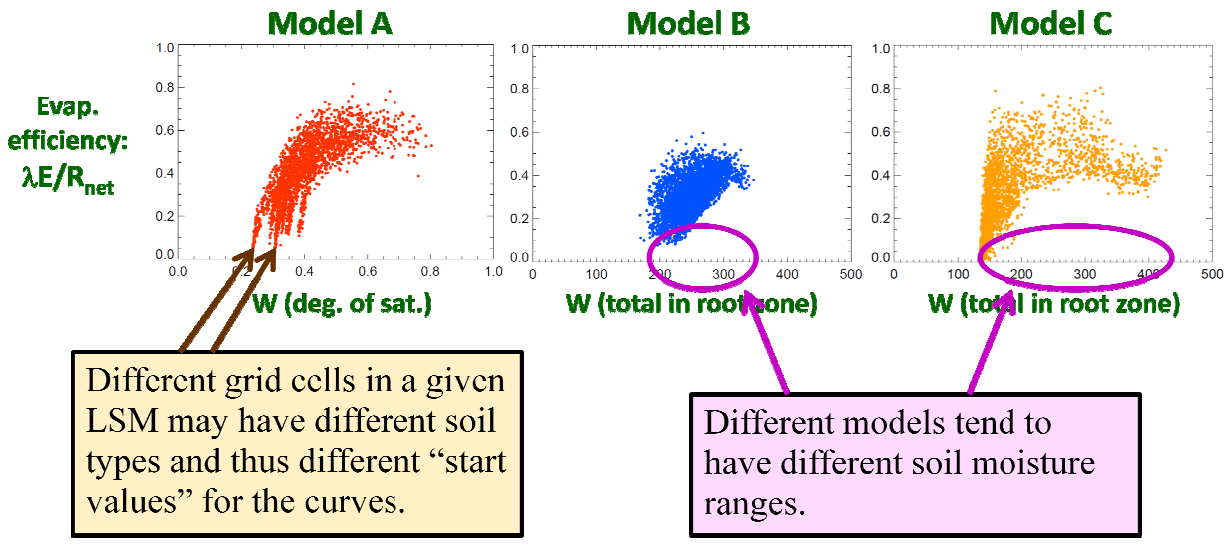
Figure 7. a. Basins (in gray) examined collectively in the analysis. See Mahanama et al. (2012) for details on the basins and on the streamflow observations taken therein. b. Color-coded efficiency space curves as in Figure 5b, but for an assumed 0.1 m soil depth and for the collected set of basins. The black curve represents the curve associated with the highest simulation skill. c. Same, but for an assumed 0.5 m soil depth. c. Same, but for an assumed 1m soil depth. d. Same, but for an assumed 1.667 m soil depth.



798

799

800 Figure 8. a. Budyko's semi-empirical relationship between the evaporation-precipitation ratio
801 and the dryness index, defined as $R_{net}/P\lambda$. b. Efficiency space curves from Figure 7c. A
802 transform of Budyko's relationship is shown in white.



803

804

805 Figure 9. Scatter plots of evaporation efficiency (ratio of monthly-averaged E to monthly-
806 averaged R_{net}) versus monthly-averaged root zone soil moisture, as derived from the May-
807 September output diagnostics produced over CONUS by three state-of-the-art LSMs.

808

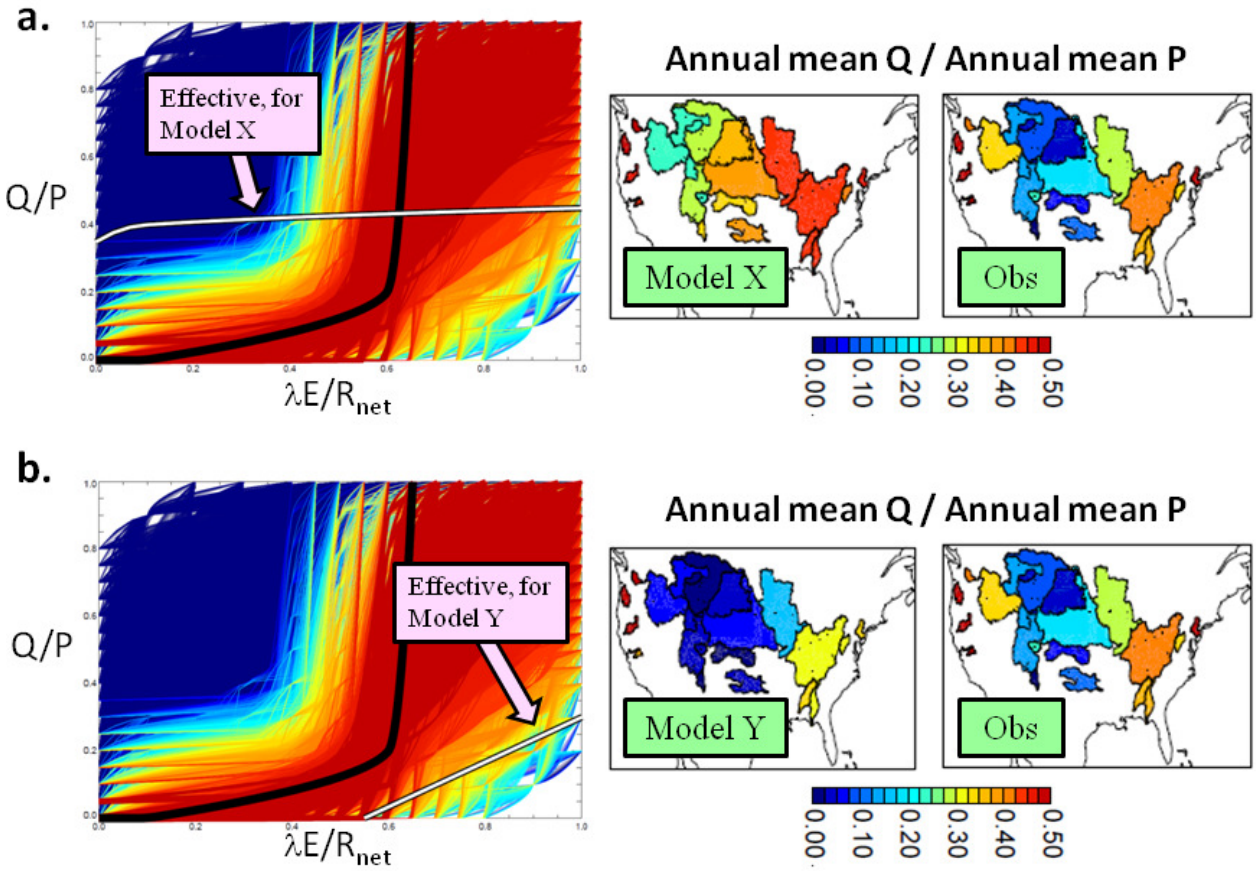


Figure 10. a. Efficiency space curve (in white) for a specific LSM ("Model X"), overlain on the color-coded curves from Figure 7c. To the right are plots of annual runoff ratio across several basins as computed from Model X diagnostics and from observations. b. Same, but for a different LSM ("Model Y").

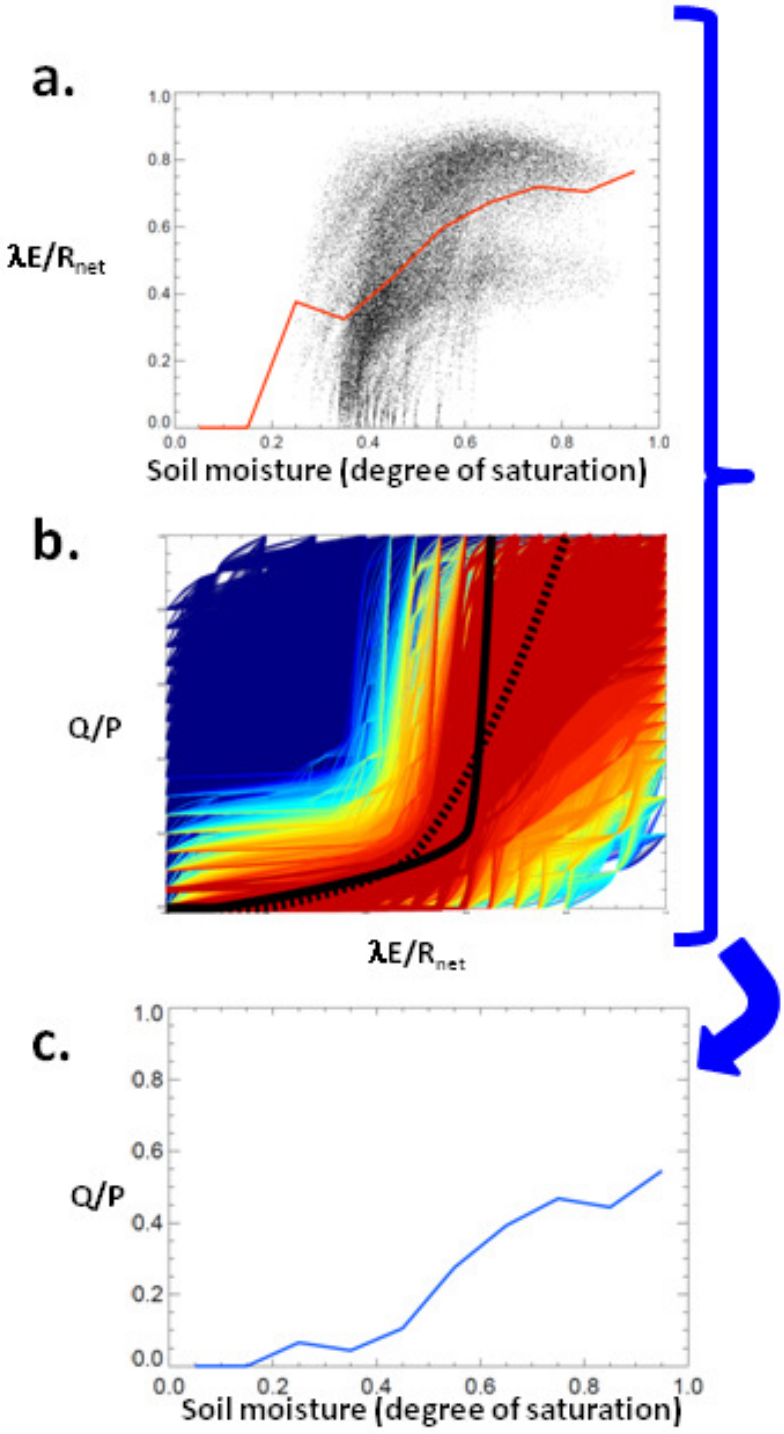
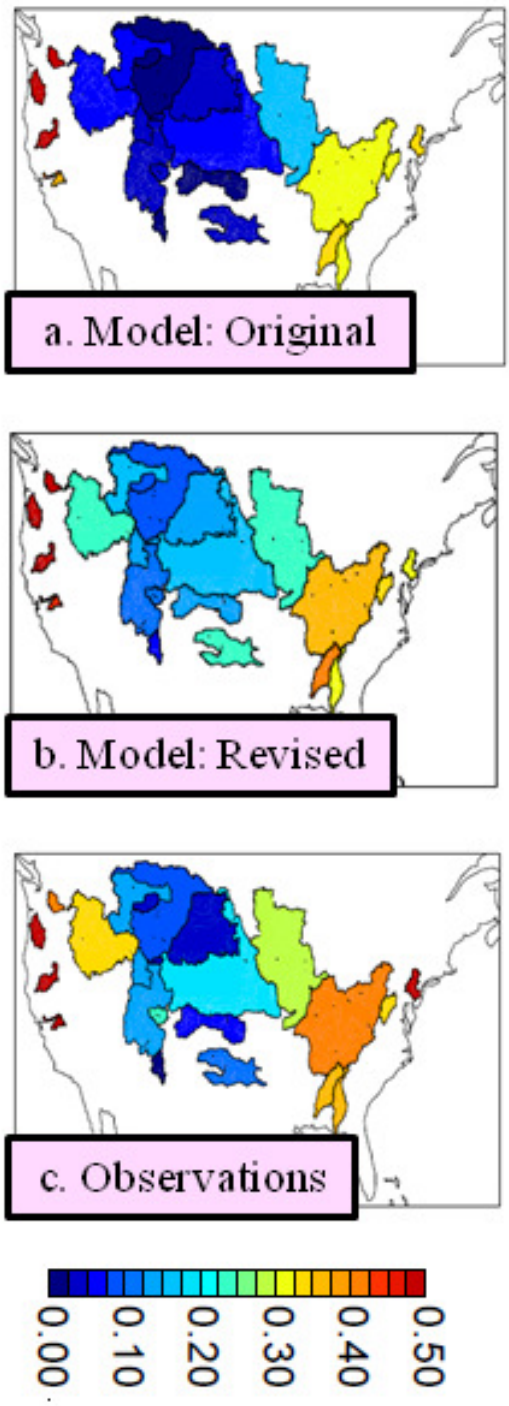
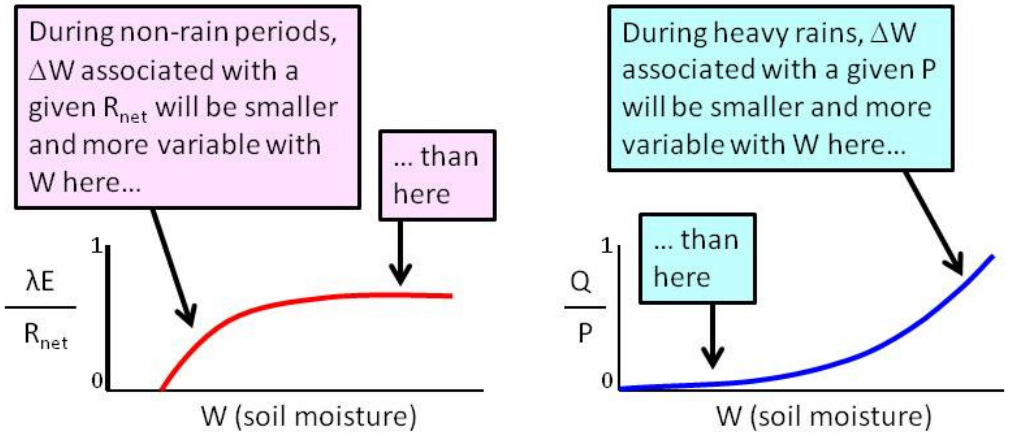


Figure 11. a. Scatter plot of evaporation efficiency (ratio of monthly-averaged E to monthly-averaged R_{net}) versus monthly-averaged root zone soil moisture, as derived from the May-September output diagnostics produced over CONUS by the Catchment-CN LSM. b. Color-coded efficiency space curves from Figure 7c, with one of the curves highlighted as a heavy dotted line. c. Resulting relationship between runoff efficiency and root zone soil moisture.



823

824 Figure 12. a. Annual runoff efficiencies (ratio of annual basin-averaged runoff to annual basin-
825 averaged precipitation) for several CONUS basins, as produced by the Catchment-CN LSM in a
826 multidecadal offline simulation. b. Same, but for a tuned version of the Catchment-CN LSM
827 (see text). c. Same, but for observations.



828

829

830 Figure 13. Sample evaporation efficiency and runoff efficiency functions, with highlighted
831 regimes of interest.